I am a doctoral candidate in the Management Department at the University of Texas at Austin (and licensed attorney; CA Bar #255347). My dissertation examines how managers evaluate and mitigate risk in making hiring, firing, and promotion decisions, and how this perception changes their behavior. Specifically, I hypothesize that managers will perceive a greater risk of litigation in the event of termination decisions, and that they will therefore (a) be more likely to discriminate against minorities in the hiring stage (this would illuminate a critical shortcoming in anti-discrimination legislation), (b) take greater steps to protect themselves before making the decision (such as documenting work history or offering termination packages), and (c) be more willing to defend those cases in court.

To test this hypothesis, I am looking to collect a nationally-representative sample of federal court cases which will then be categorized based on the type of decision that led to the lawsuit. Although the number of cases (~142,000) appears large at first glance, three factors lead me to this number. First, the three types of decisions at the center of my dissertation (hiring, firing, and termination) are not the only events that lead to litigation; for instance, sexual harassment or other forms of discrimination are included as well. Unfortunately, although these cases would have no effect on my dissertation, I am not able to tell which cases those are with existing data published by the FJC. Second, many of the cases I have examined so far have no complaints (e.g., being removed from state courts); this makes categorizing these cases difficult, if not impossible, and they will have to be omitted from the sample. And third, among the cases premised on the three employment decisions in my dissertation, failure to hire cases are rare.

To illustrate these points, I will use data I collected from the Western District of Texas (which graciously granted a fee waiver). I collected ~2,200 cases, of which 686 have been reviewed and categorized by a law student I hired as a research assistant. Of those, 17 (2.5%) were cases filed after an alleged failure to hire, while 174 (25.4%) were based on allegations of discrimination in ways not involving hiring, promotion, or termination. Moreover, 134 (19.5%) had no complaint at all. Thus, nearly 45% of the cases I have categorized so far cannot be included in my study – but I cannot tell which cases those are without the documents—and only a tiny fraction of the cases relate to one of the main decision types from my dissertation.

From a programmatic side, I am happy to structure the data collection in any way that minimizes the effect on the court. The collection will be computerized, using computer code which has already been developed and successfully used to collect the documents from the Western District of Texas. Thus, I can program how fast or slow the documents are downloaded and/or at what time of day they occur. If data collection needs to happen over the course of a month or two for districts with heavier case loads, I am certainly willing to accommodate that.

As a former practicing litigator myself, I understand the limits the courts are already under; I have no interest in adding to those worries. Rather, I need to collect a large number of case documents to ensure that I have enough data to explore the questions posed in my dissertation. I am also happy to discuss my study, terms of collection, etc., with any of the court clerks who would like further information.

Thank you!